November 27, 2006

Natural Resources GIS Program

National Park Service
U.S. Department of the Interior

NPS Data Store: Metadata and Data Uploading Guidance

Overview

The Natural Resource and GIS Metadata and Data Store application (NPS Data Store) manages and shares natural resource and GIS metadata and data generated by the Natural Resource and Servicewide GIS Programs of the National Park Service. To facilitate data dissemination to the public and throughout the National Park Service, the NPS Data Store application functions as the NPS GIS Clearinghouse and posts non-sensitive geospatial information to Geospatial One-Stop Portal (http://www.geodata.gov/gos). The NPS Data Store is part of the NPS Metadata System and provides two functions: the NR-GIS Metadata Database and the NPS Data Store Server. The NR-GIS Metadata Database is a repository of and search engine for metadata describing natural resource and GIS data. The NPS Data Store Server hosts natural resource and GIS data (documented by the metadata in the NR-GIS Metadata Database) for download.

At version 1, the NPS Data Store allows uploading of non-sensitive metadata and data. If data are posted, metadata in XML format must accompany the data file(s). Metadata alone may be posted to the NPS Data Store and the data it describes can be served from any accessible server. Also, metadata documenting data available only via offline distribution (CD/DVD, etc.) may be uploaded to the NPS Data Store.

All metadata files uploaded to the NPS Data Store become records in the NR-GIS Metadata Database. Uploaded data files are posted to the NPS Data Store Server and are linked to their associated metadata record. In addition, a copy of the XML metadata file is posted to the NPS Data Store Server. During the upload process, metadata elements may be added to the metadata record to achieve this linkage to files on the data server.

The NPS Data Store applies the following assumptions when uploading metadata and data:

- Data files must be accompanied by metadata. These files are posted together to the NPS Data Store Server.
- Data files are limited to 25Mb.
- Metadata files without associated data are posted to the NPS Data Store Server.
- If populated, the Network_Resource_Name element in the Distribution Information section of a metadata record is used to determine either:
    - where the data file is posted on the NPS Data Store Server, or
    - the data file link to data posted on other servers

  If this value is blank, the NPS Data Store will prompt the user for an upload location.
- For metadata without accompanying data, an Online_Linkage element in a cross-reference citation in the Identification Information section of a metadata record is used to determine where the metadata file is posted on the NPS Data Store Server. If this value is blank, the NPS Data Store will prompt the user for an upload location.
- For metadata records with different paths in the Network_Resource_Name and Online_Linkage elements noted above, the NPS Data Store will use the path in the Network_Resource_Name element as the upload location.

- If the filename of an existing metadata or data file on the NPS Data Store Server matches an uploaded file, the existing file will be overwritten.

At version 1, the NPS Data Store accepts single record uploads. Batch upload capability will be added in a later release.

Metadata and data upload are restricted to authorized, authenticated users of the NPS Data Store. To upload metadata and data, a user account must have Editor, Author or Author/Reader rights for the relevant NPS unit(s). To request an account, use the Account Request link at https://science1.nature.nps.gov/nrdata.

Preparing Metadata and Data for Uploading

File Naming:

- Do not use spaces in file names.
- Do not use special characters (other than underscores) in file names.
- Metadata and associated data files should have the same 'root' for their file names:
  - E.g., acadhyd.xml (metadata) and acadhyd.e00 (data)
- Verify the file names do not duplicate names of files currently in the target folder on the NPS Data Store Server (http://nrdata.nps.gov). Data files will be over-written during the upload process.

Metadata Preparation:

- Metadata must be in XML format. Text or HTML format will not be accepted.

- Ideally, metadata destined for upload should include the elements from the NPS Metadata Profile recommended in the NR-GIS Metadata Authoring Guidance and Creating Metadata (http://science.nature.nps.gov/nrdata/docs/metahelp/metahelp.cfm) documents. Use the templates for NPS Profile metadata available from the NPS Data Store Instructions help page (http://science.nature.nps.gov/nrdata/docs/metahelp/metahelp.cfm).
  If users cannot generate metadata files with these elements, the NPS Data Store will prompt for these values during the upload process.

- For geospatial data, document the spatial reference of the data in the metadata record.

- Geodatabase metadata is stored as XML in the underlying RDBMS and cannot be automatically extracted from an uploaded geodatabase file (*.mdb). For geodatabase metadata, refer to the guidance in the Posting Geodatabase Metadata section of the NPS Data Store Metadata Authoring Guidance document (http://science.nature.nps.gov/nrdata/docs/metahelp/metahelp.cfm).

- Parse metadata prior to uploading it to the NPS Data Store. Refer to the NR-GIS Metadata Parsing Guidance (http://science.nature.nps.gov/nrdata/docs/metahelp/metahelp.cfm) document for instructions.

- If using the NPS ArcCatalog© Metadata Tools Extension v1.5.1 to generate metadata, export metadata as plain XML.

- If using Metaparser to generate XML metadata files, use the NPS_Profile.cfg configuration file available in the Metadata Parsing Information section of the NPS Data Store Instructions help page (http://science.nature.nps.gov/nrdata/docs/metahelp/metahelp.cfm).

Data Preparation:

- The maximum data file size for the NPS Data Store is 25Mb. Files larger than this may be distributed on other servers or in CD/DVD format.
- For geospatial data in ESRI© Shapefile format or CAD format, zip the *.shp, *shx, and *.dbf or *.dxf components into a zipped or tarred archive file using WinZip© or a similar tool.
- For geospatial data in ESRI© export format, zip or tar the *.e00 export file(s) to reduce file transfer sizes.
- For geospatial data in ESRI© geodatabase format, zip or tar the *.mdb or SDE export file to reduce file transfer size.
- For raster data, export it to an interchange format and zip or tar the exported file. If exporting grids from ESRI© Arc/Info, apply full compression when exporting. For TIFF files, be sure to include the world file (*.tfw) in the zipped or tarred archive. For MrSID files, include the *.stw world file in the zip archive.
- For databases or spreadsheets, zip up or tar the database or spreadsheet file (*.mdb, *.dbf, *.xls).
- For executable files (*.exe), zip or tar them first since the posting process will not allow the upload of EXE files.
- XML metadata files should be included in the zip or tar archive file but at version 1, the NPS Data Store will not update metadata files in zip or tar archives.

Recommended General Practices:

- Refer to the NPS Metadata Profile (http://science.nature.nps.gov/nrdata/docs/npsprofile.cfm) for details on NPS-specific metadata elements. See the NPS Data Store Metadata Authoring Guidance document (http://science.nature.nps.gov/nrdata/docs/metahelp/metahelp.cfm) for further information.
- Use Metaparser to verify all metadata records destined for the NPS Data Store. Refer to the NR-GIS Metadata Parsing Guidance (http://science.nature.nps.gov/nrdata/docs/metahelp/metahelp.cfm) document for instructions.

Uploading Metadata and Data

Data posted to the NPS Data Store must be accompanied by metadata. This results in downloadable (online) data. In this case, the metadata record will be flagged as an 'OnlineData' record in the NPS Data Store.

If metadata was not generated with NPS Profile metadata elements, the upload utility will prompt for this information. In this situation, when data and metadata are uploaded together, the user will be prompted for:

- the NPS unit or units associated with the data
- the purpose of the metadata (e.g., OnlineData, BioProfile, etc.)
- a Theme Category derived from the NPS Theme Category thesaurus
    - E.g., Geology, Biology, Fire, etc.

- an ISO (International Standards Organization) Topic Category derived from the ISO 19115 Topic Category thesaurus
    - E.g., Geoscientific Information, Biota, Enviroment, etc.
- a general data site identifier (NR-GIS) and a data steward

Several FGDC (Federal Geographic Data Committee) metadata elements are used in the NPS Data Store. If these are missing from the metadata, the user will be prompted for:

- the title of the data set
- the point of contact for the data set (defaults to the user uploading data)
- the path for the metadata on the NPS Data Store Server
- the path for the data file on the NPS Data Store Server

If a data path is not specified in the metadata, the user must select from data paths to which they have access permissions. If more than one data file is uploaded and the folders in the data paths specified are different, the metadata file will be copied into each folder along with the data file.

After data and metadata are uploaded, the data are discoverable through the NPS Data Store search engine. In addition, the data may be browsed through the NPS Data Store Server simple browse interface (http://nrdata.nps.gov). Non-sensitive metadata records are posted periodically to the NPS GIS Clearinghouse in NPSFocus, a node of the National Spatial Data Infrastructure (NSDI).

Uploading Metadata Only

Metadata without an associated data file can be posted to the NPS Data Store. This metadata may point to downloadable (online) data housed on another server or may document an offline method for acquiring the data.

If metadata was not generated with NPS Profile metadata elements, the upload utility will prompt for this information. In this situation, when metadata only is uploaded, the user will be prompted for:

- the NPS unit or units associated with the data documented by the metadata
- the purpose of the metadata (e.g., NPS, BioProfile, etc.)
- a Theme Category derived from the NPS Theme Category thesaurus
    - E.g., Geology, Biology, Fire, etc.
- an ISO (International Standards Organization) Topic Category derived from the ISO 19115 Topic Category thesaurus
    - E.g., Geoscientific Information, Biota, Environment, etc.

Several FGDC metadata elements are used in the NPS Data Store. If these are missing from the metadata, the user will be prompted for:

- the title of the data set
- the point of contact for the data set (defaults to the user uploading data)
- the contact for the metadata

If a metadata path is not specified in the metadata, the user must select from metadata paths to which they have access permissions.

After metadata are uploaded, it is discoverable through the NPS Data Store search engine. Non-sensitive metadata records will be posted periodically to the NPS GIS Clearinghouse in NPSFocus, a node of the National Spatial Data Infrastructure (NSDI).